

WHAT IS CLAIMED IS:

1. A computerized method for designing a multi-epitope construct having multiple epitope nucleic acids, the method comprising the steps of:

5 storing a plurality of input parameters in a memory of a computer system, said input parameters comprising a plurality of epitopes, at least one motif for identifying junctional epitopes, a plurality of amino acid insertions and at least one enhancement weight value for each insertion;

10 generating a list of epitope pairs from said plurality of epitopes;

determining for each of said epitope pairs at least one optimum combination of amino acid insertions based on said at least one motif, said plurality of insertions and said at least one enhancement weight value for each insertion; and

15 identifying at least one optimum arrangement of said plurality of epitopes, wherein a respective one of said at least one optimum combination of amino acid insertions is inserted at a respective junction of two epitopes, so as to provide an optimized multi-epitope construct.

20 2. The method of claim 1 wherein said step of identifying at least one optimum arrangement comprises performing an exhaustive search wherein all permutations of arrangements of said plurality of epitopes are evaluated.

25 3. The method of claim 1 wherein said step of identifying at least one optimum arrangement comprises performing a stochastic search wherein only a subset of all permutations of arrangements of said plurality of epitopes are evaluated.

30 4. The method of claim 1 wherein said step of identifying at least one optimum arrangement comprises:

performing an exhaustive search of all permutations of arrangements of said plurality of epitopes when the number of epitopes to be included in said multi-epitope construct is less than a specified value X; and

performing a stochastic search, wherein only a subset of all permutations of arrangements of said plurality of epitopes are evaluated, when the number of epitopes to be included in said multi-epitope construct is greater than or equal to X.

5. The method of claim 1 wherein said plurality of input parameters further includes a maximum number of insertions (MaxInsertions) value, and said step of determining for each epitope pair at least one optimum combination of amino acid insertions comprises calculating a function value (F) for each possible combination of insertions for each epitope pair, wherein the number of insertions in a combination is in the range of 0 to MaxInsertions, said function value being calculated in accordance with the equation  $F = (C+N)/J$ , when  $J > 0$ , and  $F = 2(C+N)$ , when  $J = 0$ , wherein C equals the enhancement weight value of a C+1 flanking amino acid, N equals the enhancement weight value of an N-1 flanking amino acid, and J equals the number of junctional epitopes detected for each respective combination of insertions in an epitope pair based on said at least one motif.

6. A computerized method for designing a multi-epitope construct having multiple epitopes, the method comprising the steps of:

storing a plurality of input parameters in a memory of a computer system, said input parameters comprising a plurality of epitopes, at least one motif for identifying junctional epitopes, a plurality of amino acid insertions, a C+1 enhancement weight value for each insertion, a N-1 enhancement weight value for each insertion, and a maximum number of insertions (MaxInsertions);

generating a list of epitope pairs from said plurality of epitopes;

for each combination of insertions for each epitope pair, wherein the number of insertions is in the range of 0 to MaxInsertions, calculating a function value (F) using the equation  $F = (C+N)/J$ , when  $J > 0$ , and  $F = 2(C+N)$ , when  $J = 0$ , wherein C equals a C+1 enhancement weight value of a respective flanking amino acid insertion, N equals a N-1 enhancement weight value of a respective N-1 flanking amino acid insertion, and J equals the number of junctional epitopes detected for each respective combination of insertions in an epitope pair based on said at least one motif;

determining for each epitope pair at least one optimal combination of insertions yielding a maximum function value F;

generating a list of optimal combinations of insertions; and

based on said list of optimal combinations of insertions, identifying at least one optimum permutation of said multi-epitope construct comprising said plurality of epitopes arranged in an order that yields a maximum sum of function values, wherein a respective one of said optimal combinations of insertions are inserted at a respective junction of two epitopes of said multi-epitope construct.

7. The method of claim 6 further The method of claim 6 wherein said step of identifying at least one optimum permutation comprises performing an exhaustive search wherein all permutations of arrangements of said plurality of epitopes are evaluated.

5

8. The method of claim 6 wherein said step of identifying at least one optimum permutation comprises performing a stochastic search wherein only a subset of all permutations of arrangements of said plurality of epitopes are evaluated.

10

9. The method of claim 8 wherein said plurality of input parameters further comprises a maximum search time (MaxSearchTime) value and said stochastic search is performed for a period of time approximately equal to said MaxSearchTime value, wherein said at least one optimum permutation comprises at least one permutation evaluated as having a maximum sum of function values.

15

10. The method of claim 6 wherein said step of identifying at least one optimum permutation comprises:

20

performing an exhaustive search of all permutations of arrangements of said plurality of epitopes when the number of epitopes to be included in said multi-epitope construct is less than a specified value X; and

performing a stochastic search, wherein only a subset of all permutations of arrangements of said plurality of epitopes are evaluated, when the number of epitopes to be included in said multi-epitope construct is greater than or equal to X.

25

11. A computer system for designing a multi-epitope construct having multiple epitopes, the system comprising:

30

a memory for storing a plurality of input parameters, said input parameters comprising a plurality of epitopes, at least one motif for identifying junctional epitopes, a plurality of amino acid insertions and at least one enhancement weight value for each insertion;

a processor for retrieving said input parameters from said memory and generating a list of epitope pairs from said plurality of epitopes;

said processor further determining for each of said epitope pairs at least one optimum combination of amino acid insertions, based on said at least one motif, said plurality of insertions and said at least one enhancement weight value for each insertion;

5       said processor further identifying at least one optimum arrangement of said plurality of epitopes, wherein a respective one of said optimum combinations of amino acid insertions are inserted at a respective junction of two epitopes, so as to provide an optimized multi-epitope construct; and

10       a display monitor, coupled to said processor, for displaying said at least one optimum arrangement of said plurality of epitopes to a user.

12.     The system of claim 11 wherein said processor, when identifying at least one optimum arrangement of said plurality of epitopes, performs an exhaustive search wherein all permutations of arrangements of said plurality of epitopes are evaluated to identify at least one optimized multi-epitope construct.

13.     The system of claim 11 wherein said processor, when identifying at least one optimum arrangement of said plurality of epitopes, performs a stochastic search wherein only a subset of all permutations of arrangements of said plurality of epitopes are evaluated to identify at least one optimized multi-epitope construct.

14.     The system of claim 11 wherein said processor, when identifying at least one optimum arrangement of said plurality of epitopes, performs an exhaustive search of all permutations of arrangements of said plurality of epitopes when the number of epitopes to be included in said multi-epitope construct is less than a specified value X, and performs a stochastic search, wherein only a subset of all permutations of arrangements of said plurality of epitopes are evaluated, when the number of epitopes to be included in said multi-epitope construct is greater than or equal to X.

15.     The system of claim 11 wherein said plurality of input parameters further includes a maximum number of insertions (MaxInsertions) value and said processor, when determining for each epitope pair at least one optimum combination of amino acid insertions, calculates a function value (F) for each possible combination of insertions for each epitope pair, wherein the number of insertions in a combination is in the range of 0 to MaxInsertions, said function value being calculated in accordance with the equation  $F = (C+N)/J$ , when  $J >$

0, and  $F = 2(C+N)$ , when  $J = 0$ , wherein  $C$  equals the enhancement weight value of a  $C+1$  flanking amino acid,  $N$  equals the enhancement weight value of an  $N-1$  flanking amino acid, and  $J$  equals the number of junctional epitopes detected for each respective combination of insertions in an epitope pair based on said at least one motif.

5

16. A computer system for designing an optimized multi-epitope construct having multiple epitopes, the system comprising:

an input device for inputting a plurality of input parameters specified by a user;

10

a memory, coupled to the input device, for storing said plurality of input parameters, said input parameters comprising a plurality of epitopes, at least one motif for identifying junctional epitopes, a plurality of amino acid insertions, a  $C+1$  enhancement weight value for each insertion, a  $N-1$  enhancement weight value for each insertion, and a maximum number of insertions (MaxInsertions);

15

a processor for retrieving said input parameters from said memory and generating a list of epitope pairs from said plurality of epitopes;

wherein said processor, for each combination of insertions for each epitope pair, wherein the number of insertions is in the range of 0 to MaxInsertions, calculates a function value ( $F$ ) using the equation  $F = (C+N)/J$ , when  $J > 0$ , and  $F = 2(C+N)$ , when  $J = 0$ ,

20

wherein  $C$  equals a  $C+1$  enhancement weight value of a respective flanking amino acid insertion,  $N$  equals a  $N-1$  enhancement weight value of a respective  $N-1$  flanking amino acid insertion, and  $J$  equals the number of junctional epitopes detected for each respective combination of insertions in an epitope pair based on said at least one motif; and

wherein said processor further determines for each epitope pair at least one optimal combination of insertions yielding a maximum function value  $F$ , generates a list of optimal combinations of insertions, and, based on said list of optimal combinations of insertions, identifies at least one optimum permutation of said multi-epitope construct comprising said plurality of epitopes arranged in an order that yields a maximum sum of function values, wherein a respective one of said optimal combinations of insertions are inserted at a respective junction of two epitopes of said optimized multi-epitope construct.

25

30

17. A data storage device storing a computer program for designing a multi-epitope construct having multiple epitopes, the computer program, when executed by a computer system, performing a process comprising the steps of:

retrieving a plurality of input parameters from a memory of a computer system, said input parameters comprising a plurality of epitopes, at least one motif for identifying junctional epitopes, a plurality of amino acid insertions and at least one enhancement weight value for each insertion;

5                   generating a list of epitope pairs from said plurality of epitopes;

                  determining for each of said epitope pairs at least one optimum combination of amino acid insertions based on said at least one motif, said plurality of insertions and said at least one enhancement weight value for each insertion; and

                  identifying at least one optimum arrangement of said plurality of epitopes,  
10   wherein a respective one of said at least one optimum combination of amino acid insertions is inserted at a respective junction of two epitopes, so as to provide an optimized multi-epitope construct.

18.     The data storage device of claim 17 wherein said computer program, when  
15   executed, performs an exhaustive search wherein all permutations of arrangements of said plurality of epitopes are evaluated so as to identify said at least one optimum arrangement of said plurality of epitopes.

19.     The data storage device of claim 17 wherein said computer program, when  
20   executed, performs a stochastic search wherein only a subset of all permutations of arrangements of said plurality of epitopes are evaluated so as to identify said at least one optimum arrangement of said plurality of epitopes.

20.     The data storage device of claim 17 wherein said computer program, when  
25   executed, performs an exhaustive search wherein all permutations of arrangements of said plurality of epitopes are evaluated, when the number of epitopes to be included in said multi-epitope construct is less than a specified value X, and performs a stochastic search, wherein only a subset of all permutations are evaluated, when the number of epitopes to be included in said multi-epitope construct is greater than or equal to X, so as to identify said at least one  
30   optimum arrangement of said plurality of epitopes.

21.     The data storage device of claim 17 wherein said computer program, when executed, further retrieves a maximum number of insertions (MaxInsertions) value from said memory of said computer system, and further performs said step of determining for each

epitope pair at least one optimum combination of amino acid insertions comprises by calculating a function value (F) for each possible combination of insertions for each epitope pair, wherein the number of insertions in a combination is in the range of 0 to MaxInsertions, said function value being calculated in accordance with the equation  $F = (C+N)/J$ , when  $J > 0$ , and  $F = 2(C+N)$ , when  $J = 0$ , wherein C equals the enhancement weight value of a C+1 flanking amino acid, N equals the enhancement weight value of an N-1 flanking amino acid, and J equals the number of junctional epitopes detected for each respective combination of insertions in an epitope pair based on said at least one motif.

22. An apparatus for designing a multi-epitope construct having multiple epitopes, comprising:

means for storing a plurality of input parameters in a memory of a computer system, said input parameters comprising a plurality of epitopes, at least one motif for identifying junctional epitopes, a plurality of amino acid insertions and at least one enhancement weight value for each insertion;

means for generating a list of epitope pairs from said plurality of epitopes;  
means for determining for each of said epitope pairs at least one optimum combination of amino acid insertions based on said at least one motif, said plurality of insertions and said at least one enhancement weight value for each insertion; and

means for identifying at least one optimum arrangement of said plurality of epitopes, wherein a respective one of said at least one optimum combination of amino acid insertions is inserted at a respective junction of two epitopes, so as to provide an optimized multi-epitope construct.

23. The apparatus of claim 22 wherein said means for identifying at least one optimum arrangement comprises means for performing an exhaustive search wherein all permutations of arrangements of said plurality of epitopes are evaluated.

24. The apparatus of claim 22 wherein said means for identifying at least one optimum arrangement comprises means performing a stochastic search wherein only a subset of all permutations of arrangements of said plurality of epitopes are evaluated.

25. The apparatus of claim 22 wherein said means for identifying at least one optimum arrangement comprises:

means for performing an exhaustive search of all permutations of arrangements of said plurality of epitopes when the number of epitopes to be included in said multi-epitope construct is less than a specified value X; and

means for performing a stochastic search, wherein only a subset of all permutations of arrangements of said plurality of epitopes are evaluated, when the number of epitopes to be included in said multi-epitope construct is greater than or equal to X.

26. The apparatus of claim 22 wherein said plurality of input parameters further includes a maximum number of insertions (MaxInsertions) value, and said means for determining for each epitope pair at least one optimum combination of amino acid insertions comprises means for calculating a function value (F) for each possible combination of insertions for each epitope pair, wherein the number of insertions in a combination is in the range of 0 to MaxInsertions, said function value being calculated in accordance with the equation  $F = (C+N)/J$ , when  $J > 0$ , and  $F = 2(C+N)$ , when  $J = 0$ , wherein C equals the enhancement weight value of a C+1 flanking amino acid, N equals the enhancement weight value of an N-1 flanking amino acid, and J equals the number of junctional epitopes detected for each respective combination of insertions in an epitope pair based on said at least one motif.

27. A method for designing a multi-epitope construct that comprises two or more CTL epitope nucleic acids wherein the construct is presented to an HLA Class I processing pathway, the method comprising steps of:

- (i) sorting the CTL epitope nucleic acids to minimize the number of junctional epitopes;
- (ii) introducing a flanking amino acid residue selected from the group consisting of K, R, N, Q, G, A, S, C, and T at a C+1 position of a CTL epitope nucleic acids;
- (iii) introducing one or more amino acid spacer residues between two epitope nucleic acids, wherein the spacer prevents the occurrence of a CTL or HTL junctional epitope; and,
- (iv) selecting one or more multi-epitope constructs that have a minimal number of junctional epitopes, a minimal number of amino acid spacer residues, and a maximum number of K, R, N, G, A, S, C, or T at a C+1 position relative to each CTL epitope nucleic acids.



28. A method for designing a multi-epitope construct that comprises two or more HTL epitope nucleic acids wherein the construct is presented to an HLA Class II processing pathway, the method comprising steps of:

- (i) sorting said epitope nucleic acids to minimize the number of junctional epitopes;
- (ii) introducing a flanking amino acid residue selected from the group consisting of G, P, N or A positioned between said nucleic acid epitopes; and
- (iii) introducing one or more amino acid spacer residues between two epitope nucleic acids, wherein the spacer prevents the occurrence of a HTL junctional epitope.

29. The method of claim 27, wherein the spacer residues are independently selected from residues that are not known HLA Class II primary anchor residues.

30. The method of claim 27, wherein introducing the spacer residues prevents the occurrence of an HTL epitope and further, wherein a spacer comprises at least 5 amino acid residues independently selected from the group consisting of G, P, and N.

31. The method of claim 30, wherein the spacer is GPGPG.

32. The method of claim 27, wherein introducing the spacer residues prevents the occurrence of an HTL epitope and further, wherein the spacer is 1, 2, 3, 4, 5, 6, 7, or 8 amino acid residues independently selected from the group consisting of A and G.

33. The method of claim 27, wherein the flanking residue is introduced at the C+1 position of a CTL epitope.

34. The method of claim 27, wherein the flanking residue is selected from the group consisting of K, R, N, G, and A.

35. The method of claim 27, wherein the flanking residue is adjacent to the spacer amino acid residues.

36. The method of claim 27, further comprising substituting an N-terminal residue of an HLA epitope that is adjacent to a C-terminus of an HLA epitope comprised by the

multi-epitope construct with a residue selected from the group consisting of K, R, N, G, and A.

37. The method of claim 27, further comprising a step of predicting a structure of the multi-epitope construct, and further, wherein the selecting step further comprises selecting one or more multi-epitope constructs that, when introduced into a cell, is processed by an HLA processing pathway such that all of the eptiopes included in the multi-epitope construct are produced by the HLA processing pathway.

38. A multi-epitope construct prepared using the method of claim 1 or 27.

39. The multi-epitope construct of claim 37, wherein the epitopes comprised by the multi-epitope construct are encoded by a minigene.

40. A multi-epitope construct comprising a plurality of CTL epitope nucleic acids and a plurality of spacer nucleic acids, wherein:

the CTL epitope nucleic acids encode class I HLA epitopes;

the CTL epitope nucleic acids encode epitope peptides of about eight to about thirteen amino acids in length;

the spacer nucleic acids are positioned between the CTL epitope nucleic acids;

the spacer nucleic acids encode between one and eight amino acids;

one or more of the spacer nucleic acids encodes an amino acid sequence that is different than the amino acid sequence encoded by other spacer nucleic acids; and

each of the spacer nucleic acids optimizes epitope processing and minimizes junctional epitopes.

41. The multi-epitope construct of claim 40, further comprising a targeting nucleic acid.

42. The multi-epitope construct of claim 40, further comprising a nucleic acid sequence encoding a HTL-specific epitope.

43. The multi-epitope construct of claim 40, wherein two or more of the spacer nucleic acids encodes an amino acid sequence that is different than the amino acid sequence encoded by other spacer nucleic acids.

5 44. The multi-epitope construct of claim 40 having the nucleotide sequence selected from the group consisting of EP-HIV-1090, HIV-CPT, HIV-FT, HIV-TC, HCV.1, HCV.2, HCV.3s1, HCV.3s2, HCV.3s2(-3), HCV.3s3, HCV.PC3, HCV.PC4, HCV.2431(1P), HCV.4312(1P), AOSL.K, HBV.1, HBV.2, PfCTL.1, PfCTL.2, PfCTL.3, Pf33, TB.1, BCL A2 #90, BCL A2 #88, Prostate 1, and a nucleotide sequence that hybridizes to any of the  
10 foregoing.

45. The multi-epitope construct of claim 40 having the nucleotide sequence selected from the group consisting of EP-HIV-1090, HCV.3s1, HCV.3s3, HCV.PC3, HCV.PC4, HCV.2431(1P), HCV.4312(1P), HBV.2, PfCTL.1, PfCTL.2, PfCTL.3, Pf33,  
15 TB.1, BCL A2 #90, BCL A2 #88, Prostate 1, and a nucleotide sequence that hybridizes to any of the foregoing.

46. The multi-epitope construct of claim 40 having the nucleotide sequence of EP-HIV-1090 or a nucleotide sequence that hybridizes EP-HIV-1090 under stringent  
20 conditions.

47. A multi-epitope construct comprising a plurality of HTL epitope nucleic acids and a plurality of spacer nucleic acids, wherein:

the HTL epitope nucleic acids encode class II HLA epitopes of about seven to about  
25 seventeen amino acids in length;

the spacer nucleic acids are positioned between the HTL epitope nucleic acids; and

the spacer nucleic acids encode five or more amino acids, wherein each of the spacer nucleic acids optimizes epitope processing and minimizes junctional epitopes.

30 48. The multi-epitope construct of claim 47, wherein the spacer nucleic acids encode an amino acid sequence having alternating glycines and prolines

49. The multi-epitope construct of claim 48, wherein one or more spacer nucleic acids encode the amino acid GPGPG.

50. The multi-epitope construct of claim 49, wherein every spacer nucleic acid encodes the amino acid GPGPG.

5 51. The multi-epitope construct of claim 47 having the nucleotide sequence of EP-HIV-1043 or a nucleic acid that hybridizes to EP-HIV-1043 under stringent conditions.

52. The multi-epitope construct of claim 47 having the nucleotide sequence of EP-HIV-1043 PADRE or a nucleic acid that hybridizes to EP-HIV-1043 PADRE under  
10 stringent conditions.

53. The multi-epitope construct of claim 47 having the nucleotide sequence of HIV 75mer or a nucleic acid that hybridizes to HIV 75mer under stringent conditions.

15 54. The multi-epitope construct of claim 47 having the nucleotide sequence of PfHTL or a nucleic acid that hybridizes to PfHTL under stringent conditions.

55. A multi-epitope construct comprising fifteen or more epitope nucleic acids and ten or more spacer nucleic acids, wherein:

20 the epitope nucleic acids encode class I HLA epitopes or class II HLA epitopes;  
the epitope nucleic acids encode epitope peptides of about seven to about seventeen amino acids in length;

the spacer nucleic acids are positioned between the epitope nucleic acids; and  
the spacer nucleic acids encode between one and eight amino acids when inserted  
25 between the class I HLA epitope nucleic acids and five or more amino acids when inserted between the class II HLA epitope nucleic acids.

56. The multi-epitope construct of claim 55 having twenty or more epitope nucleic acid sequences and fifteen or more spacer nucleic acid sequences.

30 57. The multi-epitope construct of claim 55 having twenty-five or more epitope nucleic acid sequences and eighteen or more spacer nucleic acid sequences.